GIS NEWS AND INFORMATION

October 1996 (No. 12)

Dedicated to CDC/ATSDR scientific excellence and advancement in disease control and prevention using GIS

Selected Contents: Census 2000 and GIS-- upcoming CDC/ATSDR lecture (p.1); Stratospheric ozone and human health web site (p.2); Users provide GIS expertise to FDA (p.4) and NCHS colleagues (p.6); EPA's ENVIROFACTS digital data online (p.8); Recent public health GIS literature (p.9); PHS issues on race and ethnicity identifiers (p.13).

I. Public Health GIS (and related) Events

☛ 21st Annual Conference of the Association of Public Data Users- Addressing Critical Issues Facing the Data User Community, October 28 - 30, 1996, Holiday Inn on the Hill, Washington, D.C.; Contact: Janet Roseen, APDU96 Conference Coordinator, (612)-625-6044 or email <apdu96@mrdc.lib.umn.edu> or conference homepage at mrdc.lib.umn.edu/~apdu96.

☞ Ninth annual NCHS Geography Awareness Week guest lecture, November 21, 1996, 2:00-3:15p.m.- Bob Marx, Associate Director for Decennial Census, Bureau of the Census, "Census 2000: Implications for GIS and Survey Activities", NCHS, Hyattsville, MD. Please make **ENVISION** arrangements now for viewing off site. Contact CDC Envision personnel Eileen Dubin, Clay Cooper or Mat Fazio at (404) 639-3404.

☛ CDC and ATSDR Symposium on Statistical Methods- Statistical Bases for Public Health Decision Making: From Exploration to Modeling, Atlanta, January 28-30, 1997. Contact: Barbara McDonnell, Ph: (404) 639-3806 or email bgm4@epo.em.cdc.gov.

II. News from GIS USERS

(Please communicate directly with colleagues on any issues)

A. General News

1. From **Duane Marble**, The Ohio State University: Topic Zip Codes- Just remember that these files are subject to change without notice! I recently encountered someone matching 1996 addresses with 1990 Zip-based Census data. The individual concerned could not see why there was any problem. If you get such a database you should demand to know the metadata associated with it and what provisions are being made for timely updates. Duane F. Marble, Department of Geography, E-mail: marble.1@osu.edu, Telephone: (614) 292-2250.

2. From **Keith Clarke**, University of California Santa Barbara: The Prentice Hall Series in Geographic Information Science will have a new title in October, "Getting Started With Geographic Information Systems," by Keith C. Clarke. The book is a first freshman level text in GIS, with extensive study guide material and resource guides. The book has its own web page at www.prenhall.com/clarke. Further information about the book and a table of contents is available on the Prentice Hall web site at www.prenhall.com under Science, Engineering & Math/Science and Math/Geography/Technical Courses.

3. From **Jimmie Givens**, NCHS: On September 23, 1996, the President signed into law H.R. 3230, the "National Defense Authorization Act for Fiscal Year 1997." Among other items, the Act creates a new National Imagery and Mapping Agency, which will provide a single focus within the Government for managing imagery and geospatial information.

4. CDC's Office of Communication recently sponsored a presentation entitled "Geo-Psychographic Data Applications in Planning and Evaluating Public Health Communication Interventions", which I could

not attend. I would be interested in hearing from any GIS User who may have attended and felt there was something of potential interest to report to our user group. **Editor**

B. Technical News

5. From **Tom Arner**, EPO: I am one of the programmers of Epi Map, a product of the Epidemiology Program Office of the CDC. I am trying to locate a Summagraphics SummaSketch II digitizer with which to test our software. Any help in this matter would be greatly appreciated. **Steve Campbell**, Epidemiologist, Baltimore City Health Department: I finally got a census tract and a Zip code map of Baltimore digitized and read into Epi Map. It took a Summagraphics II digitizer and Epi Map version 1 (note: version 2 doesn't seem to work with the Summagraphics II digitizer, although it will work with a III) and a lot of calls/emails to the CDC Orkand Epi Info help desk. However, it's up and running. Steve <TQPZ46A @ prodigy.com>. P.S. By the way, once everything was set up correctly, the actual digitization process was a breeze. It was just the setup that was a major problem.

6. From **Mark Papania**, NIP: Does anyone know where I can find a map of New York City, down to Zip code areas, for Atlas or MapInfo? Please call me at (404) 639-8215.

C. Internet News

(Selected items picked up from the Internet by GIS Users)

7. From **Karen Stakes**, PHS library, NCHS (picked up from multiple recipients of list MEDLIB-L, subject- mapping web site): For a good introduction to desktop mapping and the latest news about TIGER and other data sources, visit: www.wessex.com. We offer a wide variety of mapping and demographic software and data, along with industry gossip. Our data is available in MapInfo, ArcView, SPSS, SAS, Excel and dBase formats. Samples, demos, software, manuals and utilities are available for free download.

8. From **Jon Sperling**, Bureau of the Census (Topic-CIESIN Information Stratospheric Ozone and Human Health World Wide Web Site, picked up from CIESIN User Services): CIESIN and its Socioeconomic Data and Applications Center (SEDAC) are pleased to announce the availability of the Stratospheric Ozone and Human Health World Wide Web (www) site. This is an on-line service that integrates NASA remote-sensing and atmospheric data on stratospheric ozone depletion and ultraviolet radiation with health-related data and information to provide a multidisciplinary data resource for health officials, decision makers, government officials, researchers, and the general public. The Uniform Resource Locator (URL) for the Stratospheric Ozone and Human Health www site is: sedac.ciesin.org/ozone.

The Stratospheric Ozone and Human Health www site has several components: * The Ultraviolet Interactive Service (UVIS)- This service provides ultraviolet (UV) radiation climatology data for selected locations around the United States. Average hourly, daily, and monthly estimates of biologically effective doses from 1979-1990 are accessible through an interactive query engine that allows users to visualize data of interest. These data are derived using NASA satellite data and other geophysical input parameters in an atmospheric radiative transfer model. A description of the model used to develop this service is also provided.

UVIS allows users such as epidemiologists to access detailed historical estimates of UV radiation exposure. These data can be useful for a variety of purposes, including the reconstruction of historical exposure patterns. It can help answer a range of questions such as: - how much higher are exposure levels in Albuquerque compared with Seattle? - how many days in the summer of 1988 were exposure levels higher in Detroit as compared with Atlanta? - what was the estimated cumulative DNA-damaging exposure amount during the spring of 1990 in the Salt Lake City area? - what was the range in hourly exposure levels during the month of August 1980 in Honolulu?

* Searchable Bibliographic Database: This service provides a searchable index containing more

than 3,000 citations of journal articles, conference presentations, books, and other periodicals on the topics of ozone depletion, UV radiation, and ecological and human health. Users may query the database through several search fields: text string, author, source, subject category, year, and title. Users can find results such as: - a list of publications by a particular author; - recent publications on possible links between UV exposure and skin cancer incidence; and - early efforts and studies on ecological impacts of UV exposure.

* Human Health Data Resources: This service provides a guide to statistical and epidemiological datasets and related resources from disease registries, surveys, and studies that provide information on human health effects related to UV exposure. It includes links to subnational as well as national and international data sources, disease registries, surveys, and epidemiological studies. Data sources accessible through this service include: - the New South Wales Cancer Registry; - State Cancer Registries in the United States; - the Surveillance, Epidemiology, and End Results (SEER) Cancer Statistics Review, 1973-1991; and - the North American Association of Central Cancer Registries(NAACCR).

* Related Internet Resources: This component of the Stratospheric Ozone and Human Health Www site provides a guide to additional resources available via the Internet in the areas of ozone depletion, UV radiation, and human health.

* Thematic Guide on Stratospheric Ozone Depletion and Global Environmental Change: CIESIN's Thematic Guide resource provides overview documents and full-text access to peer-reviewed publications related to stratospheric ozone depletion, potential impacts on human and ecosystem health, chlorofluorocarbons, and more.

You will need a forms-capable www browser to take full advantage of this www site. Alternately, if you have telnet access to the Internet, you may telnet to infoserver.ciesin.org and log in as "lynx" to use a character-oriented www browser. For more information, please contact CIESIN User Services by e-mail at ozone@ciesin.org or by telephone at 517/797-2727. This service is provided by the Consortium for International Earth Science Information Network (CIESIN) under contract to the U.S. National Aeronautics and Space Administration (NASA) for the development and operation of the Socioeconomic Data and Applications Center (SEDAC). SEDAC is one of the data centers in NASA's Earth Observing System Data and Information System (EOSDIS). SEDAC's mission is to develop and deliver information products and services that integrate social and natural science data in ways useful for decision making. CIESIN is a registered trademark of the Consortium for International Earth Science Information Network.

9. From **Mike Qualls**, IHPO (picked up from list SEASIA-L, subject- Symposium on population, health, environment [Thailand]): The Department of Geography, Chiang Mai University, is organizing a symposium on POPULATION, HEALTH AND THE ENVIRONMENT in Chiang Mai, Thailand, 7-11 January 1997. The symposium is being held on behalf of the IGU Commission on Population and the Environment with the participation of the Commission on Health. Pre Conference Deadlines: 1 November (registration and receipt of abstracts) and 30 November (receipt of papers). For further information about the symposium contact: Dr. Anchalee Singhanetra-Renard, PO Box 87, Chiang Mai University, Chiang Mai 50202, Thailand, Fax: +66 53 222 763, Phone: +66 53 221 699 Ext. 3572, Email: <soggi001@cmu.chiangmai.ac.th>.

10. From **Mike Mungiole**, NCHS (picked up from <gis-l@urisa.org>): (a) The Association for Geographic Information is pleased to announce the launch of the new improved dictionary of GIS terminology. The URL is: www.geo.ed.ac.uk/agidict/welcome.html. Please take the time to visit the site. Although still under review at the moment, the site is continually updated and at present contains definitions for approximately 750 GIS related terms and selected diagrams. In the near future this is expected to rise to 1500 and possibly include graphics and animation. As well as providing definitions to GIS related terminology, the dictionary

also includes terminology from Remote Sensing, Geography, Computing, Spatial Analysis, Photogrammetry and Cartography. The GIS Dictionary is brought to you by the Association for Geographic Information and the University of Edinburgh. Comments and queries regarding the GIS Dictionary may be sent to:<agidict@geo.ed.ac.uk>, or the Association for Geographic Information at <agi@geo.ed.ac.uk>. (b) Please come and visit GISVille, a new virtual GIS town complete with clickable town plan and friendly tour guide at dspace.dial.pipex.com/town/parade/fa21/gis/gis.htm. There is something for everyone at GISVille : free software - a MapBasic dialog designer and Map Key generator; official Ordnance Survey map images; examples of GIS activity and datasets in Liverpool City Council, UK; magical mystery links to related sites; academic stuff - problems of map copyright; GIS modeling etc; forthcoming attractions. Enjoy! Dave Yarwood, Liverpool City Council, UK <fa21@dial. pipex.com>.

11. **Editor**- (picked up from <gis-l@urisa.org>): (a) From: "Michael J. Kitchin" <mcoyote@clark.net>): Due to positive response and customer feedback, the availability of LabelEase--the powerful, easy-to-use label generation software for ArcView, has been extended. You may now download a fully functional, 30-day trial version of LabelEase from our web site: www.innovsys.com. You may obtain additional product information and download other GIS productivity solutions from ISD at this location as well. (b) We have a lot of PLSS data ,1:24K along the Tx/Mexican border. Jill Wolf, GIS Data Conversion Program Mgr., Infotec Development, Inc., GeoSpatial Services, 4099 SE International Way, #206, Portland, Oregon 97222, with the email address:wolf@santa.infotec.com; www.teleport.com/~idigis/GISPAGES/GeoSpatialHome.html, 503-794-1344, ext. 102. (c) From: "Darrell"@news.cais.com: Duane F Marble <dmarble@magnus.acs.ohio-state.edu> wrote ... If you are trying to generate random points (in any program) you need to be careful of the random number generator used. Most system supplied random number generators are only random in one dimension. If you should, for example, call the generator twice in a row to obtain two numbers that will be the x,y coordinates of a point you will not generally get a set of random points. On one test that I did the result was not only not random but when plotted gave something that looked like a series of parallel lines about half an inch apart and rising to the right at about 20 degrees. Response: Duane is right. Check Cressie's book on Spatial Statistics (published by John Wiley & Sons) for details on Monte Carlo testing and Skellam's statistic. It's fairly easy to do this test--as I recall, it took only an hour or two (programming included)... a small price to pay for having an answer to this often questioned procedure (random coordinate generation). I used this approach when doing dividing data into test/train sets for comparison of spatial interpolants. Using my *verified* RNG, I generated a random coordinate, found the nearest point, and put it into a test set. This way, I was assured that my test sets were as random as possible (spatially, that is). Darrell McCauley, dmccauley@casecorp.com, Senior Technical Specialist, Case Corporation. (d) The Electronic Atlas Enterprises Home Page and website is now accessible at www.electronic-atlas.com. The website offers an on-line sample of the "Electronic Atlas Newsletter," a hands-on GIS users' publication; information concerning the consulting activities of Electronic Atlas Enterprises is available as well. With respect to the newsletter, the website offers the most up to date index of the 80 issues published to date, and interested parties can order subscriptions as well.

III. GIS Outreach

(Editor: All solutions are welcome and will appear in the next edition; please note that the use of trade names and commercial sources that may appear in *GIS News and Information* is for identification only and does not imply endorsement by CDC or ATSDR)

☎ From **Kathy Godon**, FDA: I am interested in getting started in GIS applications using MapInfo and am looking at ordering software suitable for the following applications: a) Antimicrobial surveillance program which will consist of approximately 2000

samples of Salmonella, E.coli and in the near future campylobacter. Samples will originate from three types of sources with sources having varying degrees of geographic resolution, some will be linked to Zip code, county or state, b) Tissue Residue violations which will have Zip code information, and c) Use of the National Ambulatory Medical Care Survey to determine if regional differences in antimicrobial use are related to microbial resistance patterns. Questions that I have regarding the purchase of the software: 1) Map info products available include MapInfo DeskTop and MapInfo Professional, it appears the MapInfo Pro offers more display, editing and analytical capabilities than the cheaper Desk top version, is this an important feature for most users? (MapInfo Pro is a 32 bit application, MapInfo Desktop is a 16 bit application 2) Is addition of high resolution County Boundaries or High Resolution Zip Codes an option that users find is necessary or helpful with MapInfoPro? 3) Is other software compatible with MapInfo? Are there other sources of demographic and census tract information available for use with this software? 4) What other software do you suggest using with the MapInfo package? Please feel free to make any other comments and suggestions. Please send your replies to me at <Kgodon@bangate.fda.gov> , my phone number is (301)827-4224. Please include your phone number and email address.

✔RESPONSE from **Dick Hoskins**, WA State Dept. Of Health: If you have not decided to absolutely use MapInfo - I would suggest Maptitude - its a LOT cheaper, has the same functionality, actually more, free tech support, and they send you a lot of data to get started, like all you need. That saves $1000 or more right there. Also I find that its easier to learn and its faster. Its about $400-500. My shop also has MapInfo and Arcview because we have users with these packages, but Maptitude is a good alternative. I practically guarantee that you can do the same thing in this package as anywhere else. Also a 32 bit version will appear soon. They can be contacted at peter@caliper.com, this fellow has something to do with sales. I am sure you could get the software in 24 hours. Also for decent geocoding we use DataStar. Its fast and accurate (but not cheap) The geocoders in desktop software are not bad, but do not compare to DataStar. Call me if you need anymore info (360-705-6055).

✔RESPONSE from **Katherine Heck**, NCHS: I don't have answers to most of your questions, unfortunately, but I thought I could say a bit about #3. I've used Microsoft Excel with MapInfo to do mapping of disease rates by Zip code and census tract. I suspect Lotus for Windows would work with it too, but I'm not sure. It's quite easy to merge a spreadsheet with MapInfo. I believe we got some of our Zip code / census tract geocoding data from the U.S. Postal Service - and maybe some from the Census? (this was for King County, Washington, a couple of years ago) - i.e. just a list of addresses and their corresponding Zips and tracts, that we matched to, e.g., death certificate address data for mapping. Once you have the data, all you need to do is create a spreadsheet with values for the Zips (or whatever area) that matches the MapInfo file's area values, and then merge them within MapInfo. The version of MapInfo that we had came with a file that created a map of U.S. states, and we purchased smaller level maps (Zips, census tracts, and block groups) from the county. I'm not sure where else you might buy those - maybe MapInfo has some you could buy. Good luck. K.Heck, (301) 436-5979 ext 120, keh8@nch07a.em.cdc.gov.

✔RESPONSE from **Scott Hendricks**, NIOSH: I am doing a project looking at robbery rates by census tracts and am using Map-Info. I am in the process right now of converting Census-tract info into Map-Info. The Census Bureau has an amazing amount of data available on the internet. Their site is at www.census. gov. I also have their ftp sites if you are interested. Although this is free, it does take some time and programming to convert to Map-Info. Map-Info provides all the census data in its format, but it gets rather pricey. I would be willing to send you some of the programs and help you in any way I can if you end up using Map-Info. I can be reached at 304-285-6000 and my E-mail is sah5@niosr1.em.cdc.gov. I hope this answers a part of your questions.

✔RESPONSE from **Jim Wilson**, East

Carolina University: We use MapInfo Pro at ECU's Center for Health Services Research and Development. I was disappointed that county files did not come with it. You can purchase many types of data for use with MapInfo from third party vendors. It does come with a Zip code file but I found it somewhat lacking. I found better files off the web via the census bureau's Tiger Mapping Service (www.census.gov)--they have several files that can be used with MapInfo with minimal amounts of finagling: county boundary files (not real high resolution), Zip codes centroids, and smaller than county geographies e.g. census tracts (if I remember right). It is better than paying for them. Also, there is an SPSS--MapInfo link program or add-in available which makes transfer of statistical analysis results into MapInfo easy. Within Map Info you can import a variety of formats as well. Jim Wilson, Research Associate, (919) 816-2786.

☎ From **Marian MacDorman**, NCHS: Currently NCHS birth and death certificate data are not geocoded, although this is something we are looking into for the future. As yet, this research is in the very early stages, and has been on the back burner for some time due to staffing shortages, etc. We would like to learn from the Census Bureau's experience in this area. If you have any suggestions for us, we would love to hear them.

One particular problem that we have had in our vital statistics data is in identifying which events are inside and outside the city limits for major US cities. The GIS programs we have seen are primarily county and census tract-based and don't give us easy answers to the inside/outside city limits question, especially when trying to batch process 4 million records or so. Any suggestions?

✔RESPONSE from **Fred Broome**, Bureau of the Census: There are several ways to solve the problem cited below. First is to use an address geocoding system such as that in several of the GIS packages (ESRI, MapInfo, Maptitiude, etc.) and have matches return the state/county/place/MCD/IndRes codes. Then a simple lookup table (can be automated via a script in some GISs) will return the Inside/Outside and where, i.e., city state, etc. Problems are, 1. Not all GISs are equal in ability to match. 2. Place boundaries change at rate of hundred's per year. Currently about 65 to 70% of the nations households are within city style addresses and consequently within TIGER/Line '95 file coverage. Due to boundary changes, addresses along the edges of places tend to have a higher rate of mismatch (more due to annexation than to detachment). Generally this will not cause too much of a problem given the ratio of addresses inside and outside vs. the number along an edge that has changed. Hope this helps.

✔RESPONSE from **Jon Sperling**, Bureau of the Census: Geocoding addresses to census geography enables one to get the county/tract/block (smallest areal unit of census geography). Once you have the block, then you have the inside/outside problem solved since TIGER contains all legal/administrative boundaries or information to create those boundaries (i.e. block/address is in/out of city limits). I'm not very familiar with source of address on vital records (place of death, home residence) but if these are systematic and consistent, then they can be geocoded to legal entities at a specified point in time according to the address on the records (if a city style address exists). - in the future, GPS coordinates could be attached to all vital records.

☎ **Tom Richards**, PHPPO, seeks Users input on a GIS project he plans to propose. Does anyone have experience with mapping local health department (LHD) jurisdictions? Ultimately, Tom would like to create a national map of the boundaries of all local public health jurisdictions in the United States. He hopes to delineate LHD boundaries with census geography files. There are many related public health uses that this map inventory might serve. Is there any precedent for mapping LHDs? Any assistance you might provide would be appreciated. Tom can be reached at (404) 639-1944 or by email <tbr1@phppo1.em.cdc.gov>.

☎ **Cathy Cubbin**, NCHS Research Intern, seeks Users input on a GIS project she is designing: The Office of Research and Methodology at NCHS has developed a

U.S. Mortality Atlas, which includes maps of the leading causes of death by age, race, and gender for approximately 800 Health Service Areas (HSAs). HSAs are counties or groups of counties aggregated based upon where residents 65 years or older receive routine hospital care. One of the next steps in the project is to layer the death rates with other types of data, i.e. behavioral, social, economic, and environmental. I will be concentrating on developing appropriate socioeconomic measures based on the research literature, identifying and retrieving existing data sources to provide the data for these measures, and layering the data in a GIS using Arc View. I plan to begin looking at 1990 Census data at the county level and the Area Resource File to retrieve the variables and then convert to HSA. My main concern conceptually is to identify socioeconomic measures that are meaningful at the HSA level as opposed to a smaller unit of analysis such as the census tract. It may also be useful to create measures of socioeconomic inequality for HSAs based on census tract level data. I would welcome any suggestions for specific socioeconomic measures/measures of inequality or for other existing sources of data. I can be reached at (301) 436-7048, x149 or email <ckc6@nch09a.em.cdc.gov>.

☎From **Mike Fay**, ATSDR (regarding the question of the geographic distribution of telephone numbers in the last edition): I must have missed the initial query, but Section III. "Outreach" listed **Gib Parrish** asking about geographic correlation of telephone numbers. In addition to what **Owen Devine** and Mr. McGee said, I picked up some shareware a few years ago that correlated all the telephone prefixes (first three numbers) then in existence around the country with latitude and longitude. I can't remember the level of detail in terms of latlong minutes/seconds, but I can remember discerning areas around Atlanta with it, something I was working on at the time. There was logical geographical separation of Atlanta prefixes; they were not all lumped together.) A drawback is that, using the prefix, it doesn't discern individual numbers; what it actually gives you is the locale of some sort of telephone company equipment, I forget just what (its docs explain what it is, though). Whether or not this is good enough, only the user can decide.

Since it's practically free and quite easy to use, it doesn't hurt to check out. But very unfortunately, I tossed it out a few months ago since it had been years since I'd used it- wouldn't you know it!- and I can't remember the name, etc. I can only hope it could be fairly readily retrieved by an Internet shareware search, however, as I came across it on a BBS that liked to grab popular software several years ago. If it floated onto a popular BBS, it probably floated lots of other places. If somebody is really interested in the software, let me know and I may be able to dig up specifics somehow. (And we can only hope the developer keeps it up to date.) I may or may not correctly remember that NPA (for Number Plan Area) is a part of its name.

☎ **Editor**- selected items from <gis-l@urisa.org> that may be useful: (1) Subject- Tiger school districts to Atlas GIS- I am trying to extract school district boundaries from the 1994 Tiger files using Atlas GIS. I was told I can do this but I have yet to figure it out. My understanding is that the Atlas import program only imports lines and Record types 1,2,3,4,5,7 and P and therefore will not impute the school districts which are in Record type A. I do have access to PC ARC but it seems I have an older version that only coverts precensus files. A third party conversion utilities or program would also work if there is one out there. I would appreciate any help anyone can lend. I am most interested in Cook and Dupage counties in Illinois. If someone has this data and would be willing to part with it, that would be great...✓RESPONSE: Check out the Metro Chicago Information Center (MCIC). They do socio-economic GIS for the Chicago area and would probably part with the file for a reasonable price. They are listed in the Chicago phone book. Alternatively, get Dr. Doolittle from Bondata in Harrisburg PA (listed under Lisa Bontempo last time I looked). Dr. Doolittle is a very nice package for extracting boundaries from TIGER. I think I heard they were planning to add school districts to their list of extractable boundaries. Kathryn Thorne, Dept of Geography and Geology, Belknap Hall, Mansfield

University, Mansfield, PA 16933 <Kthorne@mnsfld.edu> (717) 662-4612.

(2) Subject- Digital European maps- I am looking for digital maps (country boundaries, provinces, towns, rivers, roads) of France on the Internet...✓RESPONSE: MaconUSA sells a product called "Maps and Data Professional Sets-Euro Edition" that has, amongst many other countries, detailed maps of France. Product is CD-ROM, sold in Arc Coverage, MapInfo, Idrisi, or Atlas native formats, and costs $ 1,500 US., phone (617)-254-2295.

(3) Subject- Inexpensive GIS program with links to Census data?...✓RESPONSE:Maptitude from Caliper (www.caliper.com) costs $395, runs great, and has all the data you need. You could also consider ArcView with data from Wessex.

(4) From **Arlene Siller**, NCHS: Subject- Census FIPS codes for counties in SAS format- I need to prepare a table for all of counties in the U.S.. It will be nice if I can get the county's names by the aid of FIPs code, so I don't have to type 3000 county's names. I called up SI, they told me that SAS doesn't have this function. I am wondering whether someone knows some information about it...✓RESPONSES: a) SAS doesn't have a function, but there is a template data set that comes with SAS/GRAPH from which you could create a format or whatever you'd like. The dataset is CNTYNAME. If you don't have SAS/GRAPH then you can download a file from the Bureau of the Census that you can use the same way. Let me know if you don't have SAS/GRAPH--I'll hunt up the info on the Census file and b) On our web site at www.wessex.com is a utility, available for free download, which, among other things, will give you the list you want (all county names with their FIPS). Go to the free page and download County Finder. This is a setup program. Once installed, in addition to the software, you will have a file called cntylist.txt.

IV. Special Reports

(Submissions are open to all)

A. EPA's Locational Data Improvement Project: EPA is in the process of correcting a long-standing data quality problem - improving the locations of regulated facilities, and other entities of environmental concern. At the request of EPA's Deputy Administrator Fred Hansen, and the Agency's Executive Steering Committee for Information Resources Management, the Locational Data Improvement Project (LDIP) was initiated in FY96. As a result, EPA is committed to a five-year plan for gathering latitude/longitude coordinate information and documentation for regulated facilities, pipe and stack outfalls, and monitoring and sampling locations.

The LDIP builds upon EPA's Locational Data Policy developed in 1991, and subsequent Method, Accuracy, and Description (MAD) documentation coding standards developed in 1994. Key to the success of the initiative is the recognition and involvement of State, and Local stakeholders who have gathered much of this information for use in their own environmental programs. EPA will utilize the Geographic Information Systems (GIS) Teams in its Regional Offices to act as liaisons to all interested State and Local participants. EPA also intends on involving the Federal Geographic Data Committee (FGDC) to ensure coordination throughout the Federal Government.

EPA will provide access to all new locational data and documentation through its ENVIROFACTS Data Warehouse - a World Wide Web accessible repository of information on EPA regulated facilities, including the Toxic Release Inventory (TRI), Permit Compliance System (PCS), Airs Facility Subsystem (AFS), Resource Conservation Recovery Information System (RCRIS), and Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS). In addition, the ENVIROFACTS Data Warehouse also contains spatial data and tools enabling users to create maps online or download spatial data for use in their own Geographic Information Systems. For more information contact: **Andy Battin**, EPA National GIS Program Manager, (202) 260-3061 or email <battin.andrew@epamail.epa.gov>. ENVIROFACTS can be accessed at **www.epa.gov/ngispr/**.

B. Report on the UCGIS Annual Assembly (from the

AAG GIS Specialty Group, Newsletter, 9(1), August 1996): The University Consortium for Geographic Information Science (UCGIS) is a non-profit organization of academic and research institutions committed to advancing our understanding of geographic processes and spatial and temporal relationships through improved theory, methods, technology, and data. Member institutions have the opportunity and responsibility to participate in reviewing and setting national research priorities in GIS and related specialties.

The UCGIS Annual Assembly was hosted by Ohio State University in Columbus, Ohio on 16-18 June 1996. The meeting was attended by nearly 50 delegates representing 25 of the 29 current member institutions. The primary purpose of the meeting was to develop a list of national research priorities for the period 1996-2005. These priorities are to specify GIS research contributions directly applicable to identified national needs. Prior to the meeting each UCGIS member institution was requested to submit up to five separate nominations for consideration in the final UCGIS research priorities list. Although several institutions did not make nominations, 18 members did nominate a total of 80 individual topics. Prior to the Annual Assembly, the UCGIS Research Committee, chaired by David Mark of SUNY-Buffalo, aggregated the submitted topics into 19 provisional research themes. For a listing of all submitted topics and a full description of the nominating process, please visit the UCGIS home page at: **www.ucgis.org**

Following two and one-half days of one-on-one, small group and plenary discussions and presentations, the UCGIS membership delegates voted to accept ten themes as national research priorities for the next decade. These themes are: A. Conflation/Spatial Data Integration; B. Distributed Architectures; C. Extensions to Geographic Representation; D. Geographic Cognition; E. Interoperability of Geographic Information; F. Scale G. Spatial Analysis; H. Spatial Data Infrastructure; I. Uncertainty in Spatial Data and Analyses, and; J. GIS and Society. **Editor**

C. Report on the GISDATA Specialist Meeting on 'GIS & Health' in Helsinki, Finland, 29 May - 2 June, 1996: GISDATA is a four-year scientific programme funded by the European Science Foundation, co-directed by professor Ian Masser (University of Sheffield) and Mr Francois Salge (CERCO, France), and coordinated by Dr Max Craglia (University of Sheffield). The programme has organized a series of specialist meetings in three research clusters; database design, data integration, socio-economic and environmental applications. The meeting in Helsinki was the fourth of the last cluster. The meeting was organized by professor Anthony C. Gatrell (Lancaster University) and professor Markku Loytonen (University of Turku, Finland), and supervised by Max Craglia. The organizers gave the meeting three objectives; compilation of a research agenda, encouragement of cross-national, multidisciplinary research, and publication of the edited essays presented at the meeting. Paper were presented by 15 speakers representing geography, statistics, epidemiology, environmental sciences, public health, and biosciences. Papers were divided into three categories and presented in joint sessions followed by vivid discussions. The first category included papers concerned with the future of GIS and health research, organizational questions, and spatial statistics. The second group consisted of paper dealing with the provision of health care and with data related issues. The third category dealt with applications and case studies.

Below is a complete list of all papers presented. They will be published as edited essays in a book 'GIS and Health' (Gatrell & Loytonen eds.) by Francis & Taylor 1997. Nanja van den Berg, The Development of an Epidemiological Spatial Information System in the Region of Vorpommern, Germany; Mario Braga, Cesare Cislaghi, Giorgio Luppi & Carola Tasco, A Multipurpose Interactive Geographic and Epidemiological Atlas; Sue Collins, Modeling Spatial Variations in Air Quality Using GIS; Anthony C. Gatrell & Markku Loytonen, GIS and Health Research in Europe: A Position Paper; Yves Guermond, Michel Bussi & Frederic Bizet, Psychotropic Drugs Consumption in an Urban Area; Robert Haining, Spatial Statistics and the Analysis of

Health Data; Geoffrey M. Jacquez, GIS, Health and the Scientific Method; Martin Kuldorf, Statistical Methods for Spatial Epidemiology: Tests for Randomness; Gonzalo Lopez-Abente, Bayesian Analysis of Emerging Neoplasms in Spain; Markku Loytonen, GIS, Time Geography, and Health; Andrew Lovett, Robin Haynes, Graham Bentham, Sue Gale, Julii Barnad & Gisela Suennenberg, Improving Health Needs Assessment Using Patient Register Information in a GIS; Gerard Rushton, Improving the Geographic Basis of Health Surveillance Using GIS; Henk J. Scholten, Development of a Health and Environment GIS for the European Region; LyLy Teppo, Problems and Possibilities in the Use of Cancer Data by GIS - Experience in Finland; Stefania Trinca, GIS Applications to Environmental and Health Related Issues in Italy; Andre van der Veen, Rainer Fehr & Klaus Prator, Organisation and Tools to Exchange Spatial Data for Environmental Health Information Management (EHIM); Paul Wilkinson, Geographical Analysis in Health Planning. Submitted by Anthony C. Gatrell and **Markku Loytonen**.

V. Public Health GIS Literature

(This section may include literature citations, abstracts, syntheses, etc., and submissions are open to all)

❒ Aral, MM., **Maslia, ML.**, Ulirsch, GV., and **Reyes, JJ.** (1996), Estimating exposure to volatile organic compounds from municipal water-supply systems: Use of a better computational model, *Archives of Environmental Health*, (51) 4, July/August, pp. 300-309. **Abstract**: The Southington, Connecticut, water-supply system is characterized by a distribution network that contains more than 1,700 pipeline segments of varying diameters and construction materials, more than 186 mi (299 km) of pipe, 9 groundwater extraction wells capable of pumping more than 4,700 gal/min (0.2965 m^3/s), and 3 municipal reservoirs. Volatile organic compounds, which contaminated the underlying groundwater reservoir during the 1970s, contaminated the water-supply system and exposed the town's residents to volatile organic chemicals. We applied a computational model to the water-supply system to characterize and quantify the distribution of volatile, organic compounds in the pipelines, from which we estimated the demographic distribution of potential exposure to the town's residents. Based on results from modeling analyses, we concluded the following: (a) exposure to volatile organic compound contamination may vary significantly from one census block to another, even when these census blocks are adjacent to each other within a specified radius; (b) maximum spatial spread of contamination in a water-distribution system may not occur under peak demand conditions, and, therefore, maximum spatial distribution of exposed population also may not correspond to peak demand conditions, and (c) use of the proposed computational model allows for a more refined and rigorous methodology with which to estimate census-block-level contamination for exposure assessment and epidemiologic investigations.

The following papers appear in *Statistics in Medicine*, Symposium on Small Area Statistics in Public Health: Design, Analysis, Graphic and Spatial Methods, S. Jay Smith (Guest Editor), Vol. 15, No. 17/18, September, 1996:

❒ Carvalho, MS., Cruz, OG. and Nobre, FF. Spatial partitioning using multivariate cluster analysis and a contiguity algorithm, pp. 1885-1894. This paper describes an approach for combining small geographic units to stabilize mortality rates by pooling information across areas according to specified risk profiles. The procedure is based on a principal component analysis, followed by a cluster analysis of socio-economic indicators to classify the risk profile of each area. The classification is used in an algorithm to join neighboring areas with similar profiles until an estimated population size is achieved. We applied this method to two Administrative Regions of the city of Rio de Janeiro, Brazil, using the census tracts as the basic areal unit. Census tracts were classified according to four socioeconomic categories distributed spatially as a mosaic, where tracts of differing categories neighbor each other. The aggregation algorithm produced a new partition of the region studied, with the created areal units preserving the internal socioeconomic homogeneity.

❒ **Croner, CM., Sperling, J.** and **Broome, FR.**

Geographic Information Systems (GIS): New perspectives in understanding human health and environmental relationships, pp. 1961-1977. **Abstract**: Geographic Information Systems (GIS) and digital computer technology will advance the mission of the Centers for Disease Control and Prevention (CDC) and Agency for Toxic Substances and Disease Registry (ATSDR) to protect public health. Geographic positioning, topology, and planar and surface measurements are basic GIS properties which enable highly precise locational referencing of spatial phenomena. The growing uses of remotely sensed imagery and satellite facilitated Global Positioning Systems are contributing to unprecedented surveillance of the environment and greater understanding of known and suspected environmental disease associations with human and animal health. Earth science and public health monitoring GIS databases offer new analytic opportunities for disease assessment and prevention.

❑ **Kirby RS.** Toward congruence between theory and practice in small area analysis and local public health data, pp. 1959-1866. **Abstract**: Advances in computer hardware, software and database interfaces have provided opportunities for collaboration, manipulation, analysis and display of spatial data on an unprecedented scale. Demands for small area data in public health, fueled in part by an increasing emphasis on benchmarking in relation to Year 2000 objectives but also in response to state and federal programmes to involve local communities in the assessment and planning process have simultaneously generated an unprecedented demand for these analyses and data presentations. This paper discusses four areas where geographic, cartographic and statistical theory and methodology need to be brought to bear on the development of applications involving small area health data. These areas are: (i) the theoretical conception of space; (ii) managing the inherent variability of rates and frequencies; (iii) attribution of events or cases to areas or to points; and, (iv) the application of sound principles of cartographic design to the presentation of results.

❑ **Rushton, G.**, Krishnamurthy, R., Krishnamurti, Lolonis, P. and Song, H. The spatial relationship between infant mortality and birth defect rates in a U.S. city, pp. 1907-1919. **Abstract**: The spatial patterns of infant mortality and birth defect rates in the Des Moines, Iowa, urban region are described as a contoured surface based on the punctual kriging of address-matched vital statistics records from The Iowa Department of Public Health. Areas defined as having high rates are shown to be sensitive to the size of the spatial filtering units. There is no correlation between infant mortality and birth defect rates in the region. The significance of areas with high rates is determined by a Monte Carlo simulation procedure. One area of high infant mortality is found in the region, which contrasts with many smaller areas with high birth defect rates in the region. The observed birth defect rate pattern is consistent with the hypothesis that each birth was equally likely to result in a birth defect, while the infant mortality pattern is unlikely to be the result of such an equal likelihood process.

❒ Wittie, PS., **Drane, W.** and Aldrich, TE. Conventional methods fail to provide adequate assignments of individuals into appropriate subgroups for small area analyses, especially when studies cover only a portion of a county or use ecologic, descriptive variables. An unconventional aid to epidemiologic investigations, such as a geographic information system (GIS), can aid public health investigators in determining denominators for studies of disease where overlays of environmental exposure and public health data define the geographic extent of the population being studied or where data must be analyzed using a map. The basic functions of a GIS- to store, retrieve, transform and display data which have real earth-based co-ordinates- can facilitate this type of ecologic approach. In this paper, each denominator is chosen using the aid of a GIS to allow overlays of environmental elements with residential areas.

Other Literature

❑ Corbley, KP. (1996). Epidemiologists track virus-bearing mosquitoes, *GIS WORLD*, 9(3), March, pp. 52-57. Epidemiologists estimate that some 25,000

persons (mostly children) die annually in the Western Hemisphere from malaria or the bite of the female *Anopheles* mosquito. Resistance to pesticides appears to be increasing the mosquito-borne diseases worldwide of malaria, yellow fever and dengue. Vegetation and tree classes are delineated to classify Landsat Thematic Mapper imagery that can be correlated to hydrologic and topographic maps.

❒ Lee, H. (1966). Health care of San Francisco's homeless, *Geo Info Systems*, 6(6), June, pp. 46-47. An estimated 5-6,000 persons are homeless in San Francisco, lacking health insurance and thus access to most traditional fee-for-service providers, HMOs, and hospitals other than for emergency care exception being San Francisco General Hospital). Buffer zone analysis (clinic and hospital emergency room locations, shelter and free-eats (soup kitchens) sites, bus routes and the homeless by Zip code) is used to determine accessibility to health care services.

❑ **Ruiz, M.** (1996). Health statistics mapping software, *Geo Info Systems*, 6(6), June, pp. 52-55. Three BioMedware products used for the visual display of statistical clustering of health events are reviewed. The review includes Stat!, for spacial clustering (Moran's I, Grimson's method, etc.), temporal clustering and time series (Scan, Grimson's method, Larsen's method, Cuzick and Edward's method, etc.) and space-time interaction (Knox, Mantel, k-Nearest Neighbor, Grimson's method, etc.); C2D, for creating correlograms or a procedure to evaluate spatial patterns of data where each observation is a known distance from other observations (measured intervals or spatial lags), and; Geodat, a quick and easy way to create and plot maps which could also serve as a digitizer.

❒ Tobias, RA., Rima, R., Alo, CJ. and Howe, HL. (1996). Tracking human health statistics in "RadiumCity", *Geo Info Systems*, 6(7), July, pp. 50-53. GIS was used in a project designed to produce health outcomes information for the city of Ottawa, Illinois, which the federal government has declared a Superfund site because of excess radiation from 14 unrestricted sites. ATSDR is interested in health statistics related to selected NPL (National Priority List) sites. Radiation exposures are important in public health since there exists scientific links between radium-226 and radon-222 and a variety of cancers. The risks to developing fetuses and young children are also high. A variety of epidemiologic and spatial analytic techniques were employed in the study: census track and block group standardized cancer incidence ratios (SIRs), risk exposure ring zones (similarly using SIRs), nearest neighbor analysis (clustering) and Poisson Chi Square mapping (to detect association between radiation areas and disease incidence).

VI. Related Census, DHHS and Other Developments

Excerpts from the March 12-13, 1996 meeting of the NATIONAL COMMITTEE ON VITAL AND HEALTH STATISTICS, Public Health Service, Washington, D.C.: **UPDATE ON DEPARTMENT DATA ACTIVITIES**

Mr. Ebeler [Deputy Assistant Secretary in the Office of the Assistant Secretary for Planning and Evaluation] said he would focus on the DHHS Data Council's plans for the future, and he referred to a document outlining the long-term agenda. Work plans are now being developed, which will be shared with the Committee. He invited the Committee's help in shaping the Council's agenda. The long-term plan has six major themes, some built around Council tasks and some involving its role as a forum. They are as follows:

1. Develop a Department-wide data collection strategy, including coordination and integration of surveys and oversight of surveys and general statistical analysis. This has been the initial focus of the Data Council, centering on the Department's survey integration plan. Integrating the Department's research and evaluation strategy with data collection (which Dr. Raub discusses below) is another dimension.

2. Coordinate HHS and Inter-Department health data standards activities. This involves the

Department's participation in private sector standards development activities, an area in which the Department knows it needs to "get up to speed."

3. Serve as HHS liaison for the National Committee on Vital and Health Statistics. It is hoped that NCVHS will help the Data Council stay on task and bring external perspectives and interests to bear on its operations.

4. Serve as focus for HHS issues relating to privacy of health and social services information. A work group has been set up on this topic.

5. Provide a forum for coordination of health and human services issues raised by the expanding National Information Infrastructure (NII) activities.

6. Provide a forum for coordination of HHS responses to external requests for HHS action on issues related to health and social services data.

Many of the Data Council's current activities relate to these six themes. Survey integration is a major priority, and is being implemented. Another major activity is getting a new charter and new members for NCVHS. Nominations are due by March 22, and Mr. Ebeler urged the Committee to make known its desires about both specific individuals and categories of people it would like to see represented.

UPDATE ON LEE REPORT ON SES DIFFERENTIALS IN HEALTH

Dr. Madans (NCHS) said she is appearing as a representative of Dr. Lee, to give an update on the report on SES differentials in health and to try and clarify some previous misunderstandings stemming from briefings to the Subcommittee on Health Statistics for Minority and Other Special Populations. It has been decided, in consultation with Dr. Lee, that, because they have taken on different objectives, the two major initiatives that were at first combined as the Lee Report on SES Differentials in Health will be separated. One component is a chart book comprised of about 30 charts on aspects of socioeconomic differentials in health, which will be designed primarily for policy makers. The other component is a group of in-depth analytical pieces. The chart book will be done under the direction of Drs, Jack Feldman and Diane Makuc in the NCHS Office of Analysis, Epidemiology and Health Promotion, the group that is responsible for *Health U.S.* Both of them have done considerable research on SES differentials. It will take about a year to produce the chart book, which will be called the Chart Book on SES Differentials, developed under the sponsorship of the Assistant Secretary. NCHS will be looking for the most appropriate vehicle for the analytic original research pieces, which will go into the literature either as a special journal issue, as individual papers, or as another kind of combined publication.

Turning to a new subject, Dr. Madans referred to previous discussions with the Committee about translating the National Health Interview Survey. She said NCHS is exploring with the Census Bureau the capabilities of its bilingual interviewers, and also trying to build more flexibility into the HIS. If these and other matters can be resolved, then it plans to translate the questionnaire once it is finalized, sometime in 1997. It would be too expensive to translate the current version, which is expected to undergo changes during the implementation phase. Dr. Williams praised the Center's efforts in this area, and supported the idea of waiting until the questionnaire is finalized before translating it. Dr Madans added that contrary to previous reports, the language in which the interview was conducted will be included on the tape.

<u>Excerpts from the June 4, 1996 meeting of the National Committee on Vital and Health Statistics, SUBCOMMITTEE ON HEALTH STATISTICS FOR MINORITY AND OTHER SPECIAL POPULATIONS, Public Health Service, Washington, D.C.</u>

Dr. Williams [Chair] welcomed those present and noted that this meeting is a continuation of themes the Subcommittee has been following for many months regarding race and ethnicity identifiers on databases. This presentation on managed care was arranged after the Subcommittee was informed last September that one-fourth of HMO enrollees come from racial minority populations, raising concerns about how they are identified and about the quality of the data for minority Medicare and Medicaid managed care enrollees.

RACE AND ETHNICITY DATA FOR MEDICARE AND MEDICAID MANAGED CARE POPULATIONS

Mr. Clark began by describing the current status of Medicare and Medicaid programs and the trends toward more managed care. Eleven percent of Medicare beneficiaries are now in managed care. HCFA contracts with 289 HMOs, of which 202 are risk HMOs. The enrollment growth in 1995 was 26 percent, and 20 percent in 1994. Seventy-four percent of Medicare beneficiaries have access to at least one HMO setting, with the greatest density in the West.

For Medicaid, 11.6 million beneficiaries (32 percent) are now in managed care, representing a growth of 50 percent from 1994 to 1995 and another 63 percent the previous year. Forty-nine states operate a managed care program of some sort, most for the AFDC population. It is predicted that the bulk of managed care enrollment will shift to the disabled and chronically ill populations in institutional settings. They now represent 28 percent of the beneficiaries but 70 percent of the spending.

HCFA is just beginning to implement the Medicare choice demonstration, an attempt to stimulate penetration into markets with low managed care organization (MCO) participation. There has been a basic shift in HCFA's thinking from just being a payer, modeled after the insurance industry, to being a purchaser of care. With that shift has come an interest in quality and accountability, pursued through outcomes measures.

Turning to the subject of managed care for minority and special populations, Mr. Clark noted that market incentives now reward the withholding of care. Another concern is managed care's impact on providers, both because new economic forces make it difficult for traditional community providers to compete, and because many minority providers have been excluded from managed care plans. As a result, fundamental community networks and services are being disrupted, and minority and special populations are the most vulnerable to underutilization of services. Another HCFA concern is market abuse, and it has developed guidelines for plans to use in marketing their programs so as not to mislead prospective enrollees. Changes in emergency room use are another focus of concern. The Office of Managed Care has a beneficiary access and education team, which advocates for beneficiaries and tries to improve services and to assure access.

Regarding data, Mr. Clark noted HCFA's efforts to upgrade its enrollment file with race and ethnicity data and to clarify the categories of Other, Asian and Hispanic. It is developing measures for evaluating access to care, and preparing to begin collecting encounter data. There is a debate now in the agency about what data should be included and how they should be used. A statewide pilot test of encounter data is underway in Washington State. HCFA also hopes to test its access measures and beneficiary education approaches, and to develop and disseminate culturally sensitive materials. The Office is planning a survey of 4 million Medicare members to measure access, satisfaction and quality, with a telephone follow-up.

Mr. Washington commented that the federally qualified HMOs are doing a great deal on their own to get information on minority members and others with special needs. Also, HCFA is starting a strong emphasis on getting more HMO service in rural areas.

Asked about HCFA's risk adjustment methodology, Mr. Washington said the agency's expertise is limited in this area, and it is counting on being able to get input from experts to develop that information. The plans document financial and quality issues, and nothing regarding race and ethnicity at this point. Some plans, however, do it on their own. The HEDIS-type requirements ask plans to develop geographically sensitive information. Mr. Washington stressed that this is a collaborative effort among all parties, and HCFA wants to be "particularly nonintrusive in the HMOs."

Dr. Williams posed what he called a "fundamental question": given that current market forces and incentives would disadvantage people with higher medical needs--adversely affecting minority populations in particular--and given the government's historic role as a counterbalance to the marketplace, what is HCFA doing to create safeguards and counter the perverse incentives? Mr. Washington said the Office of Operations is "particularly sensitive" to that. It requires that service areas be inclusive and that

plans offer standard services, and it works with state Divisions of Insurance, who know "where poor people are."

Mr. Weil [Office for Civil Rights] expressed disappointment that the presenters had not provided more focused information on HCFA's efforts regarding race and ethnicity issues in the context of managed care. He urged that race and ethnicity be collected at every encounter because discrimination can happen at every encounter, but both Mr. Forrey of ASTM and Dr. Williams affirmed that it is preferable to collect and store this information in an enrollment data base and then to be able to link it to encounter data. Dr. Williams added that this is the National Committee's position, with the proviso that it be updated periodically because self-reported racial status does change over time.

Ms. Perot referred to Title VI and asked if HCFA requires managed care organizations to engage persons of different ethnic groups, and if information on that is collected systematically and routinely. Expanding on the point, Dr. Williams commented that research has shown that the quality of health services and outcomes is linked to the race of the provider, making it an important variable to look at. Mr. Washington said that HCFA does not require it, but he supports such a measure.

REPORT BY DR. JAMES THOMPSON

Dr. Thompson [subcommittee member] said he hoped the Subcommittee would help look into this matter, which concerns the contracts for Indian tribes' taking over their health care from the Indian Health Service. These contracts exempt the tribes from any federal confidentiality regulations regarding health care data. He explained that many regulations have been removed from the contracts to minimize the burden on the tribes, but removal of this particular one is not fair to Indian people. He asked if the Subcommittee could gather more facts on it and possibly present it to the National Committee for possible action.

UPDATE ON REVIEW OF FEDERAL STANDARDS FOR RACE AND ETHNICITY DATA

Mr. Tucker [BLS] said that as part of its review of Directive 15 concerning racial and ethnic classification standards, OMB convened an interagency committee and asked for research on five issues: the effect of including a multiracial category; the effect of combining Hispanic origin and race into a single question; current understandings of the concepts of race, ethnicity and ancestry; terminological preferences; and the possibility of new classifications such as Arab-American. Mr. Tucker and Roderick Harrison co-chair the research working group, whose first step was to conduct a supplement to the Current Population Survey on race and ethnicity.

The supplement was done in May 1995 to some 60,000 households, with questions on the first three research issues listed above. The survey used four panels with different test questionnaires, assigned randomly to 15,000 households each, to gauge the effects of variations in the questions. Mr. Tucker described the variations among these panels in the wording and/or order of the questions. The nonresponse in the labor force section of the CPS was 6.5 percent, and an additional 10.6 percent did not respond to the supplement. The responses were weighted for nonresponse, and comparisons were made across treatment conditions. The results were as follows:

•In panels with a separate question for Hispanic ethnicity, between 10 and 11 percent of the population identified as Hispanic. This dropped to 8 percent in panels that treated "Hispanic" as a racial category. When an Hispanic racial category was included, fewer people identified themselves as white.

•The inclusion of a multiracial category resulted in 1.6 percent of the population identifying as multiracial. This had little or no effect on blacks or Asian/Pacific Islanders, but resulted in a drop of about 20 percent in the American Indian population.

•In all panels, the majority of Hispanics said they would like Hispanic to be included as a racial category.

•Regarding terminology, 62 percent of whites preferred the term "white." 57.8 percent of Hispanics preferred the term "Hispanic," compared to 11.7 percent for "Latino." Among blacks, 44 percent preferred "black" and 40 percent preferred "African

American" or "Afro-American." More than half of American Indians and Alaska Natives favored those terms, while about a third prefer the more generic term "Native American." Age and national origin have an effect on preferences. For example, Mexican Americans prefer the term "Latino."

The research also compared the supplement responses to those to the original CPS questions. It found consistency among whites and blacks in regard to race and a high degree of inconsistency among American Indians. When Hispanic is treated as a race, there is 80 percent consistency for that group.

From these findings Mr. Tucker drew the following conclusions in addition to those implied above:

•The size of the Hispanic population will be larger if there is a separate question. It would drop if Hispanic were treated as a race, and the size of the white population would be smaller by 5 percent. Mexican Americans will make up a larger proportion of Hispanics if Hispanic is included as a racial category.

•The American Indian and Alaska Native population would probably be reduced if a multiracial category were included.

•A large portion of the "Other" race category will move to Hispanic if that is available as a racial category.

•Cubans can be expected to behave differently from other Hispanics on both race questions.

The BLS will soon release a final report on this survey. It will also do an evaluation for OMB to help them make decisions about race and ethnicity classifications. In addition, the Census Bureau is now carrying out several different studies in anticipation of the 2000 census.

In response to a question, Mr. Tucker noted that the term "ethnicity" is difficult for people conceptually, so using it is problematic in large-scale population surveys.

Mr. Forrey described the ASTM position in regard to race and ethnicity on computer-based patient records. The organization defines race as a biological attribute and ethnicity as a sociologic attribute, and it stresses the need to "carefully define those terms and give them different labels." This comment generated a lively discussion, particularly on the nature of race and what can, or cannot, be definitively said about it. Several people observed that there is no scientific consensus on what a race is, and Mr. Tucker noted that Directive 15 states that both race and ethnicity are treated as cultural, political or social categories for the government's purposes. Dr. Williams explained that no known biological criteria can be used to unambiguously identify a person's race. Furthermore, there is more genetic biological variation within any given racial group than between them. He suggested that the group leave this discussion with an agreement that these constructs are fuzzy.

[CDC and Office of Public Health Race and Ethnicity Project]

David Cantor then reported briefly on a project Westat is conducting for CDC and the Office of Public Health to look at the implications of changing the way race and ethnicity are asked on the birth certificate -- specifically, evaluating how multiracial mothers interpret race and ethnicity on the birth certificate. They plan to test alternative items for the race item that incorporate options to provide multiracial status. They also are analyzing the thought and response processes of multiracial women when they answer either a standard race item or one that includes multiracial status.

The project is recruiting about 700 mothers with at least one child age 0-3. Five hundred of these are multiracial, meaning that they define their parents as being of two different races. The most successful recruitment method for this quite small population proved to be through direct mail using a list of new mothers. Those who wanted to participate called on an 800 line and were screened. Qualitative cognitive interviews have now been conducted with about 40 women, and a mail survey is in progress. It will be followed by a telephone interview, with another set of race questions to test different wording and order. The race questions being compared include 1) an open-ended question that asks about race and includes examples, 2) the same question with Multiracial included as one of the examples, and 3) the same with a "Mark all that apply" option.

The researchers also are asking for authorization from the mothers to get the original birth record of their two youngest children, to compare survey results with what was written on the birth certificates. This will be done in 10 states.

So far, the cognitive interviews have been analyzed. One finding is that for multiracial persons, the standard race items are very confusing. The method by which such people choose race varies considerably, and many women choose different races on different forms, depending on the context. Most multiracial women interviewed prefer a provision for putting the specific racial combinations they identify with. Mr. Cantor said these were difficult interviews in that the subject is difficult for respondents to talk about.

[Census Bureau Projects]

Dr. Williams thanked Mr. Cantor, and invited Roderick Harrison from the Census Bureau to report. Mr. Harrison distributed summaries of two current Bureau projects, which he also described. The National Content Survey was fielded in March, and data are now being processed, with file tabulation expected to begin in August. It is a sample of 60,000 households and is being used to test a multiracial category similar to that tested on the CPS, plus some sequencing of the race and Hispanic origin variables, and alternative terminology. In June, a race and ethnic targeted test will be conducted in 114,000 households, targeting areas with high concentrations of white ethnic groups, blacks, American Indians, Asians, and Hispanics. Questions and variations similar to those in the other surveys will be tested.

Regarding the study of funeral directors' behavior, Dr. Carter-Pokras [OMH] reported that it is being studied because of widespread underreporting of the race/ethnicity of the decedent, especially for American Indians, Alaska natives, Pacific Islanders and Hispanics. This has serious implications for the quality of mortality statistics. Dr. Robert Hahn will be asked to brief the Subcommittee on his study of the way funeral directors establish race.

STATE MULTIRACIAL LEGISLATION

Dr. Carter-Pokras then reported on state activities relating to the multiracial category, acknowledging the help of Dr. Tom Levine and Nora Jimenson in preparing the information. She noted that the increase in interracial marriages in the U.S. has created a controversial issue for OMB in that multiracial families feel it is unfair to require them to identify with a single racial category. Also, researchers note that the nonresponse to the standard race and ethnicity categories and the use of the Other category have increased. Responding to multi-racial advocacy groups rather than waiting for OMB to make a decision about this, three states (Georgia, Indiana and Michigan) have passed, and ten are considering, laws requiring a multiracial category.

This legislation is taking health departments by surprise. The three aforementioned states require the use of a stand-alone multiracial category in all state forms, while some state laws only affect school forms and applications; however, the latter may be expanded to apply to health forms.

Since this legislation is relatively recent, the extent of the change it will bring about is not known. Dr. Carter-Pokras explained that OMB's directive on reporting and presenting data on race and ethnicity by federal and state governments and other organizations was issued in 1977 and has come under increasing criticism. Most data systems do not allow individuals to report more than one race. In such cases, the options for multiracial persons are to select either the category they identify with most closely, the mother's category, or the father's category. The first option is the most common, although NCHS uses the second for natality data. The Center has found that no simple algorithm can be used to assign the race of children in interracial families.

Dr. Carter-Pokras mentioned the practices of several data systems in respect to multiracial persons, anticipating some of the options now being considered by OMB. The multiracial advocacy group Project Race has requested that OMB consider a multiracial category followed by a "check all that apply" approach. However, all existing state legislation asks only for a stand-alone multiracial category, without the attendant detail. Every state also has a reallocation

rule by which, if the federal government does not have a multiracial category, the multiracial state group will be reassigned on the basis of the racial and ethnic distribution of the rest of the state population. NCHS has asked that this reallocation not be done for birth certificate data. The Decennial Census shows that the self-identification of multiracial persons does not correspond to the race and ethnicity distribution of the population.

Other problems with the state legislation include the state-to-state variations in the definition of a multiracial person and in the coverage (i.e., what documents the law applies to), as well as in legal interpretations. Also, the preferred terminology for multiracial individuals is not clear cut. Cost is another issue: Georgia reports that the change has not been costly for its vital records system, but Michigan is experiencing significant costs.

The Office of Minority Health is concerned that a stand-alone multiracial category without information on the racial background of the individual makes health risk assessment impossible. Also, since people tend to use the terms "race" and "ethnicity" interchangeably, those with mixed ethnicity may mistakenly identify with the "multiracial" category. This has implications for Hispanics and American Indians. Another concern is that the multiracial category will draw people from each of the current categories. The Indian Health Service and the Bureau of Indian Affairs expect an adverse impact on their counts, and thus on their funding -- possibly as much as a 25 percent drop. The changes in state practices also have implications for federal statistics, in that the numerators (e.g., for number of births) will not match the denominators.

Dr. Carter-Pokras reiterated that the major issue is that because the health status of multiracial children varies by the race of the mother, it is important to have details about the racial backgrounds of individuals who classify themselves as multiracial. The OMH recommends specific questions on this rather than a "check all that apply" follow-up question. Decisions are also needed on how to code and tabulate multiracial persons.

Asked to comment on the reasons for adding a multiracial category, Dr. Carter-Pokras and Dr. Williams said legislators have been moved by compelling stories about the psychological damage done to children who are forced to identify with only one of their parents. Dr. Williams noted that an OMB criterion for race and ethnicity categories is respect for individual dignity. He observed, however, that the arguments about damage have been made without good empirical scientific evidence that it in fact occurs.

Dr. Thompson observed that an effort should be made to educate American Indians and Alaska Natives that they would "shoot themselves in the foot" if they chose the multiracial stand-alone category, because of the resulting cuts in IHS funding. Mr. Harrison said some tribes have expressed concern to the Census Bureau about this.

Dr. Williams noted that the trend toward identifying with multiracial status, which may grow considerably, could destroy all of the nation's racial categories. Dr. Carter-Pokras stated that the OMH position is that multiracial people should be able to identify themselves as they perceive themselves, but that this should not be accomplished at the expense of having more detailed information on the person's origins. Also, OMH is opposed to the reallocation clauses as currently written.

Mr. Albores said that many civil rights advocates are concerned about the multiracial category, especially in the Census 2000, because the categories were created in the first place to combat discrimination. Others seconded this statement, including a representative of the HHS Office for Civil Rights.

ASIAN/PACIFIC ISLANDER SUMMIT DATA RECOMMENDATIONS

Dr.[Moon] Chen [Ohio State U.] acknowledged the "central role" of this Subcommittee in monitoring the health status of the population. He prefaced his presentation with a review of the context for the two aforementioned meetings.

Asian Americans and Pacific Islanders are the fastest growing of four federally defined minority groups, having increased by 108 percent between 1970 and 1980 and by 142 percent in the previous decade.

The Census Bureau predicts that by the year 2050, Asian/Pacific Islanders will represent 11 percent of the U.S. population. The growth is due both to immigration and to the breadth of the geographic area assigned to this category. It extends from Mongolia in the north, Indonesia in the south, Japan in the west and the Middle East in the west, plus the Pacific Islands. "This," Dr. Chen noted, "is diversity." He illustrated the diversity with the fact that India alone has 15 languages on its currency and encompasses more than 800 dialects.

Roughly two-thirds of all Asian Americans are foreign born, compared with 8 percent of the U.S. at large. Two-thirds of all Asian/Pacific Islanders speak a language other than English in the home. Acculturation is an important variable with this population, because of the marked changes between Asian/Pacific Islander immigrants and their descendants.

On the subject of data gaps, Dr. Chen noted that very few of the more than 300 Healthy People 2000 objectives target Asian/Pacific Islanders, because of the absence of baseline data which were a precondition for objectives. In addition, a Medline search of 1992 citations showed over 300 on the health status of Blacks and only 48 on the health status of Asian Americans and Pacific Islanders. The lack of scholarship and data leads many people to the mistaken conclusion that this population group has few health problems. This is a reason for the formation of the new health journal, Asian American and Pacific Islander Journal of Health [Dr. Chen, Editor]. A third data gap cited by Dr. Chen is the extent of disparity between race/ethnicity classification at birth and at death. For whites, the misclassification rate was 1.2 percent; for Filipinos (the Asian/Pacific Islander group with the highest misclassification rate), it was 79 percent. Research is another problem area: although NIH is funding more research targeted at racial and ethnic groups, few grants target Asian Americans or Pacific Islanders.

These conditions led the federal government, through the U.S. Public Health Service, to fund two major conferences on the health status of Asian/Pacific Islanders. CDC took the leadership in gathering a coalition of federal agencies which collectively funded the first National Health Summit of Asian American and Pacific Islander Organizational Leaders. This was held in June 1995 in San Francisco, with more than 200 delegates there representing 11 time zones, 26 states, and 14 ethnic groups. That conference produced 76 agency-specific recommendations. The second meeting, the Pacific Islander/Asian American Biomedical Research Conference, was held in January 1996 in Honolulu, Hawaii.

Both meetings concluded that data are needed. Dr. Chen organized their recommendations in terms of the construction, collection and commitment required tocollect the data. In terms of construction, they recommended that the data be ethnically specific, longitudinal, comparative, and inclusive of cultural factors. In terms of collection, the meetings recommended that the approaches overcome linguistic and cultural barriers (e.g., the pervasive mistrust of government) as well as the usual "national mentality" of federal data collection. Dr. Chen noted that Asian/Pacific Islander people live in concentrations around the country, and data collection can be more efficient if it takes that into account.

In view of all of these factors, the specific recommendation is for the employment of bilingual, bicultural, indigenous staff in a three-year study focusing in the first year on specific ethnic groups in six or seven selected cities (e.g., Japanese and Native Hawaiians in Honolulu, Korean Americans in Los Angeles, Filipino Americans in San Francisco, and so on). Proposed for year two is a study of Pacific peoples in Pacific jurisdictions (e.g., Guam, American Samoa, etc.), with a component to train people for capacity building. Data collection would be modeled after what the NCI has done with the SEER Program. Evaluation would be done in year three.

Dr. Chen said there is empirical evidence that the proposed indigenous model of data collection works. He cited research in Ohio through an NIH grant, which had a response rate of over 87 percent. What is needed to make this work, he said, is a commitment to cultural competence and overcoming cultural barriers.

Dr. Williams asked Dr. Chen for specific

suggestions for Subcommittee actions in support of the summits' recommendations. He noted that contrary to stereotypes, some Asian groups have the highest poverty levels in the United States. He offered several suggestions for Subcommittee recommendations--e.g., an Asian/Pacific Islander NHANES survey similar to that done for the Hispanic population. Another suggestion was requesting an analysis of several years' worth of NHIS data to get a picture of Asian/Pacific Islander health. It was noted that NCHS plans to do this, and the Subcommittee might write a letter supporting this initiative. Dr. Williams noted that in addition to the work being done on the content of the NHIS in respect to Asian/Pacific Islanders, work is needed on sampling. Dr. Carter-Pokras noted that another possibility is to work directly with states. OMH would welcome a Subcommittee statement of support for its work with the states.

Dr. Wan [subcommittee member] noted that more information is needed on aging in respect to Asian/Pacific Islanders, many of whom live alone and in deprived situations, making long-term care a matter of great concern.

Dr. Parsons [NCHS] observed that many people see the need for an organizational entity focused specifically on minority health data issues, to provide a formal institutional mechanism for coordinating efforts to look at minority health issues. She noted that the Department is and will be dependent on the ability to draw on the expertise of people like Dr. Chen. Dr. Williams noted that besides data gaps, another source of inadequate problem definition is the limitations of those sitting at the table. He suggested that the Subcommittee recommend that the appropriate expertise be brought into the relevant places. Dr. Schwartz observed that this applies to the entire Department, and in fact to the entire federal government. Dr. Carter-Pokras noted that OMH has a Resource Persons Network and DHHS has a Directory of Minority Health and Human Services Data Resources. The latter is on the HHS home page.

On the language issue, Mr. Albores [NAPALC] called attention to the proposed congressional legislation that would make English the official language of the U.S. government, thus foiling efforts to translate surveys and other forms. This bill has a possibility of passing the Senate Governmental Affairs Committee, where it is scheduled for markup in mid-June.

Final Thought(s)

It's hard to get something by **Russ Kirby**, UW-Madison Medical School. His response to the note in the last edition regarding temperature rounding error follows: As for body temperature, it's true . . . 98.6 degrees F rarely appears as the mean, median or mode in a randomly selected series of humans. For that matter, do you think that there are any randomized, controlled clinical trials to support use of penicillin for 7 days for some therapy, 10 for others, and 14 for yet others? Not much!!

From **Steve Botman**, NCHS: During the heat of the space race in the 1960's, the U.S. National Aeronautics and Space Administration (NASA) decided it needed a ball point pen to write in the zero gravity confines of its space capsules. After considerable research and development, the Astronaut Pen was developed at a cost of one million U.S. dollars. The pen worked and also enjoyed some modest success as a novelty item back here on earth...the Soviet Union, faced with the same problem, used a pencil.

Charles M. Croner, PhD, Editor, ***GIS NEWS AND INFORMATION***, Office of Research and Methodology, National Center for Health Statistics <cmc2@nch09a.em.cdc.gov>

Time for Fall pumpkins...and stay in GIS touch